Cognitive Processing Style Analyzer: Using Language to Predict Cognition

RIT Rochester Institute of Technology

RIT | AWARE-AI NSF Research Traineeship Program

Department of Psychology- Matthew Altobelli (mga4278@rit.edu), Dr. Tina Sutton (tmsgsh@rit.edu), Dr. Cecilia Alm (coagla@rit.edu)

Introduction

Goal:

The goal of this algorithm is to take novel impact statements or narratives and assess the implicit cognitive processing style of the writer.

- Three cognitive processing styles: accommodation, overaccommodation, assimilation.
- Schema related

Research Questions:

- RQ1: What symptoms of trauma, depression, and anxiety correlate with the different cognitive processing styles?
- RQ2: Are there signals in the language used to describe traumatic events that can predict cognitive processing styles?
- RQ3: If there are reliable signals, can we use those to create a machine-learning algorithm to predict processing styles from novel trauma narratives?

What is the current procedure?

- Annotations are done by hand, usually between multiple annotators.

KEY
Accommodated
Overaccommodated
Assimilated

A woman was on her phone and therefore hit someone driving towards them. It made me mad. She was distracted and could have easily gotten someone killed.	A woman was on her phone and therefore hit someone driving towards them. It made me mad. She was distracted and could have easily gotten someone killed.	A woman was on her phone and therefore hit someone driving towards them. It made me mad. She was distracted and could have easily gotten someone killed.

Figure 1. Annotations by three different raters of the same narrative. All raters annotated this statement differently.

Overall Agreement[a,b]

	Kappa	Asymptotic Standard Error	Asymptotic z	Asymptotic Sig.	Asymptotic 95% Confidence Interval Lower Bound	Asymptotic 95% Confidence Interval Upper Bound
Overall Agreement	.742	.020	37.228	.000	.703	.781

a. Sample data contains 345 effective subjects and 3 raters.
b. Rating category values are case sensitive.

Table 1. Inter-rater reliability between three raters on all annotations.

Procedure

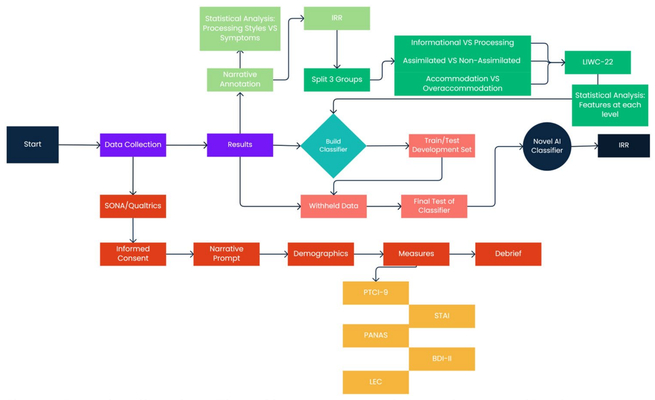

Figure 2. Procedure flow chart. The red boxes represent the experiment used for data collection. Yellow boxes are the measures. Light green boxes represent analysis prior to the hierarchy split. Dark green boxes represent information stored after the hierarchy split. The aqua box represents building the classifier. The pink boxes represent training and testing. The dark blue box represents the new model, and the black box is the final inter-rater reliability.

Results

RQ1: Correlation Results

Sample 1	Sample 2	N	Correlation 95% CI for ρ	P-Value
NEG WORLD SCORE	Percent Accommodated	97	-0.238 (-0.417, -0.040)	0.019
Word Count	Percent Assimilated	97	-0.268 (-0.444, -0.072)	0.008
PANAS NEGATIVE SCORE	Percent Informational	97	-0.225 (-0.406, -0.027)	0.027
Word Count	Percent Informational	97	0.266 (0.071, 0.442)	0.008
BDI SCORE	Percent Overaccommodated	97	0.424 (0.246, 0.575)	0
NEG SELF SCORE	Percent Overaccommodated	97	0.399 (0.216, 0.554)	0
PANAS NEGATIVE SCORE	Percent Overaccommodated	97	0.268 (0.073, 0.444)	0.008
PANAS POSITIVE SCORE	Percent Overaccommodated	97	-0.25 (-0.428, -0.053)	0.014
STAI STATE SCORE	Percent Overaccommodated	97	0.237 (0.039, 0.416)	0.02
STAI TRAIT SCORE	Percent Overaccommodated	97	0.396 (0.213, 0.552)	0

Table 2. Pearson's pairwise correlation table. Red are negative correlations, and green are positive.

- Symptoms of depression, anxiety, and posttraumatic stress show up in written language when someone is writing about a traumatic event.
- The more balanced a person's thinking is, the less negative their views about the world.
- Overaccommodated thinking is associated with higher scores for depression, negative self-score, negative affect score, and higher state and trait anxiety.

RQ3: Machine Learning Model

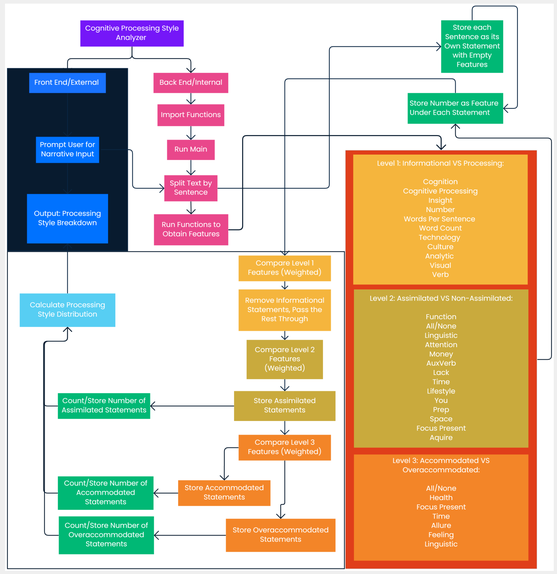

Figure 5. Flow chart of the cognitive processing style analyzer pseudocode and model. Items in the black box are what the user sees. Purple boxes represent the main backend procedures. Green boxes represent data storage. The red box represents the functions used for the linguistic features. The light orange box is level 1. The gold box is level 2. The dark orange box is level 3. The white box is the machine learning predictions based on the feature data. The light blue box represents the final calculations.

RQ2: Important Language Features

	%Accom	%Overaccom	%Assim
%Overaccomodation	-0.135		
%Assimilated	-0.25	-0.24	
%Informational	-0.396	-0.303	-0.586

Table 3. Pearson's pairwise correlation table of processing styles.

- The correlation matrix of processing styles shows the justification for separating the processing styles into a three-tier hierarchy.

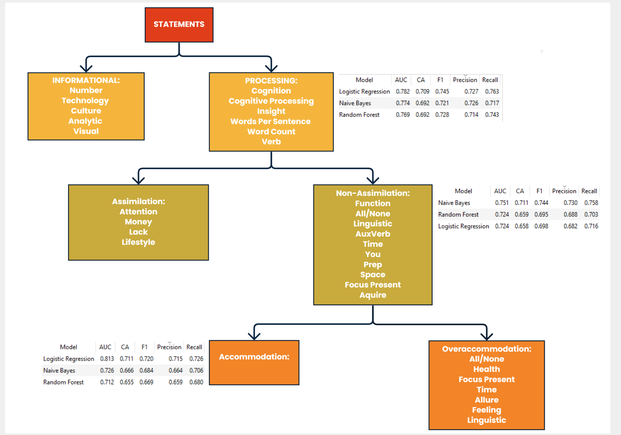

Figure 3. Hierarchical breakdown of processing styles by feature importance. Features that are significantly larger are placed in the processing-style boxes. Each level is tested by three classifiers: Naïve Bayes, Logistic Regression, and Random Forrest.

- Models with these features were tested using Orange, a data mining software (figure 4).
- There are important features of language that can predict processing styles.
- We can use these features to build a classifier that could take novel statements and generate processing style scores from them.

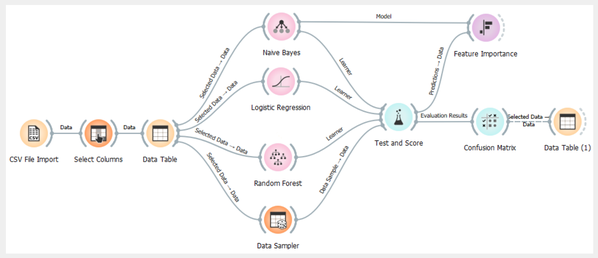

Figure 4. Orange data mining flow of classifiers.

Discussion and Conclusion

- Results suggest that there are features of depression, anxiety, and post traumatic stress that show up in the language used to describe adverse events.
- The language used to describe these events can be used as a window into the writer's cognition.
- Cognitive processing styles are the lens through which we view the world, so understanding a patient's processing style is critical for therapist practicing cognitive processing therapy.
- This analyzer would reduce subjectivity in annotations by only relying on the frequencies of words in the important features dictionaries.
- It would standardize the annotation process and produce comparable results across research studies.
- It would also save time and money in research by reducing the need for additional raters who would need to be trained.

References

American Psychiatric Association. (2017). *Diagnostic and statistical manual of mental disorders: Dsm-5*.

Balota, D.A., Yap, M.J., Cortese, M.J., Hutchison, K.A., Kessler, B., Loftis, B., Neely, J.H., Nelson, D.L., Simpson, G.B., & Treiman, R. (2007). The English Lexicon Project. Behavior Research Methods, 39, 445-459.

Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Manual for the Beck Depression Inventory–II*. San Antonio, TX: Psychological Corporation.

Boyd, R. L., Ashokkumar, A., Seraj, S., & Pennebaker, J. W. (2022). The development and psychometric properties of LIWC-22. Austin, TX: University of Texas at Austin. https://www.liwc.app

Demsar J, Curk T, Erjavec A, Gorup C, Hocevar T, Milutinovic M, Mozina M, Polajnar M, Toplak M, Staric A, Stajdohar M, Umek L, Zagar L, Zbontar J, Zitnik M, Zupan B (2013) Orange: Data Mining Toolbox in Python, Journal of Machine Learning Research 14(Aug): 2349−2353.

Dondanville, K. A., Blankenship, A. E., Molino, A., Resick, P. A., Wachen, J. S., Mintz, J., Yarvis, J.S., Litz, B.T., Borah, E.V., Roache, J. D., Young-McCaughan, S., Hembree, E.A., & Peterson, A. L. (2016). Qualitative examination of cognitive change during PTSD treatment for active duty service members. *Behaviour Research and Therapy, 79*, 1–6. https://doi.org/10.1016/j.brat.2016.01.003

Philippi, C. L., Dahl, G., Jany, M., and Bruce, S. E. (2019). Impact statement coding of self- related thought in women with posttraumatic stress disorder. Journal of Traumatic Stress, 32(2):269–276.

Sobel, A. A., Resick, P. A., & Rabalais, A. E. (2009). The effect of cognitive processing therapy on cognitions: Impact statement coding. *Journal of Traumatic Stress*, *22*(3), 205–211. https://doi.org/10.1002/jts.20408

Spielberger, C. D. (1983). State-trait anxiety inventory for adults.

Sweeting, H. and West, P. (1994). Life events checklist–revised.

Watson, D., Clark, L. A., and Tellegen, A. (1988). Positive and negative affect schedule.

Wells, S. Y., Morland, L. A., Torres, E. M., Kloezeman, K., Mackintosh, M.-A., & Aarons, G. A. (2017). The development of a brief version of the Posttraumatic Cognitions Inventory (PTCI-9). *Assessment*, *26*(2), 193–208. https://doi.org/10.1177/1073191116685401

Acknowledgement

This material is based upon work supported by the National Science Foundation under Award No. DGE-2125362. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.